

# Interoperability of Digital Libraries

Report on the work of the  
EC working group on DL interoperability

Dr. Stefan Gradmann  
Universität Hamburg / Regionales Rechenzentrum  
[stefan.gradmann@rrz.uni-hamburg.de](mailto:stefan.gradmann@rrz.uni-hamburg.de)  
[www.rrz.uni-hamburg.de/RRZ/S.Gradmann](http://www.rrz.uni-hamburg.de/RRZ/S.Gradmann)

- Working Group **context**
- Working Group **methodology**
- Evaluation of selected **interoperability frameworks**
- **10 DL Interoperability Short Term Agenda Issues**
- Long Term **Strategy Elements**
- Interoperability WG and **The European Digital Library** / EDLnet

## ***Interoperability WG Context + Mission***

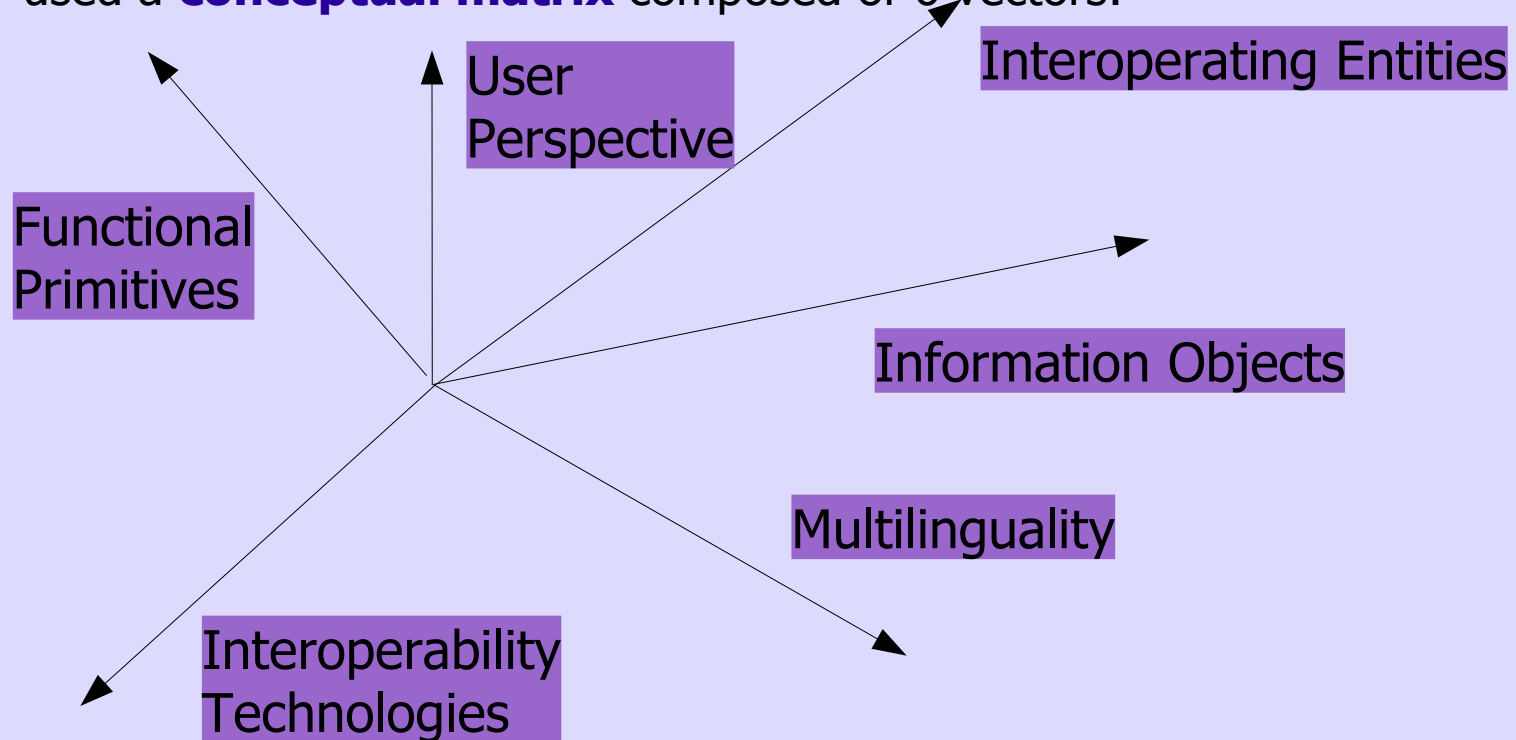
- EC i2010 (**Lisbon!**) agenda with Digital Libraries as one of 3 'flagship initiatives': the setting up of the **European Digital Library** as a common multilingual access point to Europe's distributed digital cultural heritage including all types of cultural heritage institutions
  - **2008**: at least 2 million digital objects; multilingual; searchable and usable; work towards including archives.
  - **2010**: at least 6 million digital objects; including also museums and private initiatives.
  - "I am **not** suggesting that the Commission creates a single library. I envisage a **network** of many digital libraries – in different institutions, across Europe." V. Reding (29 September 2005)
- WG active from January to June 2007 with a double mission:
  - Contribute to the **short term DL agenda** => identify areas for short term action and recommend elements of an action plan (**list of prioritised feasible options**)
  - Contribute to the **long term DL agenda** => identify key elements for a **long term strategy**

## Working Group Composition

- Emmanuelle **Bermes** (Bibliothèque nationale de France / F),
- Mathieu **Le Brun** (Centre Virtuel de la Connaissance sur l'Europe / LU)
- Sally **Chambers** (The European Library Office / TEL),
- Robina **Clayphan** (The British Library / GB),
- Birte **Christensen-Dalsgaard** (State and University Library Aarhus / DK),
- David **Dawson** (The Museums, Libraries and Archives Council / GB),
- Stefan **Gradmann** (Hamburg University Computing Center / D),
- Stefanos **Kollias** (Technical University of Athens / GR),
- Maria Luisa **Sanchez** (Ministerio de Cultura / ES),
- Guus **Schreiber** (Vrije Universiteit Amsterdam / NL),
- Olivier **de Solan** (Direction des Archives de France / F)
- Theo **van Veen** (Koninklijke Bibliotheek / NL)
- *EC: Pat **Manson** Chair), Marius **Snyders** (European Commission, DG INFSO, Cultural Heritage and Technology Enhanced Learning) Federico **Milani** (European Commission, DG INFSO, eContentPlus)*

# Conceptual Framework of Interoperability WG

- “Interoperability is the capability to **communicate, execute programs, or transfer data** among various **functional units** in a manner that requires **minimal knowledge** of the **unique characteristics** of those units.”
- To identify more precisely the determining factors of interoperability we used a **conceptual matrix** composed of 6 vectors:



### ■ Interoperating Entities

- Cultural Heritage Institutions (libraries, museums, archives)
- Digital Libraries,
- Repositories (institutional and other),
- eScience/eLearning platforms or simply
- 'Services'

### ■ Objects of Inter-Operation

- full content of digital information objects (analogue vs. born digital),
- representations (librarian or other metadata sets),
- surrogates,
- functions,
- services

### ■ **Functional Perspective of Interoperation**

- Exchange and/or propagation of digital content (OA/Non OA)
- Aggregation of objects into a common content layer (push vs. harvesting / pull)
- interaction with multiple Digital Libraries via unified interfaces
- operations across federated autonomous Digital Libraries (such as searching or meta-analysis for e. g. impact evaluation)
- common service architecture and/or common service definitions or aim at building common portal services.

### ■ **Multilinguality**

- Multilingual / localised interfaces,
- Multilingual Object Space (dynamic query translation, dynamic translation of metadata or dynamic localisation of digital content)

### ■ Design and Use Perspective

- manager,
- administrator,
- end user as consumer or
- end user as provider of content,
- content aggregator,
- a meta user or a
- policy maker.

### ■ Interoperability Enabling Technology

- Z39.50 / SRU+SRW
- harvesting methods based on OAI-PMH
- web service based approaches (SOAP/UDDI)
- Java based API defined in JCR (JSR 170/283)



# *Interoperability Abstraction Levels*

Abstract



## semantic

allowing to access similar classes of objects and services across multiple sites, with multilinguality of content as one specific aspect

## Interoperability Group Focus

## functional / pragmatic

based on a common set of functional primitives  
or on a common set of service definitions

## syntactic

allowing the interchange of metadata and protocol elements

## technical/basic

common tools, interfaces and infrastructure  
providing uniformity for navigation and access

Concrete

## *Interoperability Frameworks Discussed*

- **DELOS framework for DLs**
- **5S framework**
- **DRIVER repository federating architecture**
- **Object Reuse and Exchange (ORE)**
- **JISC Information Environment (SOA)**
- **JCR (Java API)**
- Deliberately discarded: DAREnet, aDORe. CORDRA / IMS DRI (CP and ECL), e-Framework, O.K.I. Open Service Interface Definitions (OSIDs), and more ...

# *Short Term Agenda Issues for 2008 / 1*

## ■ **(1) User Requirements**

Existing use cases in operation with TEL and the BNF 'maquette' to be used as input for a systematic and generalised process for identifying EuDL user requirements.

## ■ **(2) Object Models (granularity and structure)**

Granularity and complexity of the digital information objects will be at the level of **complete objects**. E. g. 'Books' and 'Articles' (librarian), 'records' and 'files' (archival) and 'artefacts' (museum) objects. For the longer term this should be further refined to a model for granularity that can deal with intra-object reference structures.

## ■ **(3) Persistent Identifiers**

It should be **technically impossible to create new resources in EuDL without applying standard identifiers**. Any of the known identifier frameworks (URN, DOI, ARK and others) may be used as long as they are applied systematically and the resolving mechanisms are transparent. The CENL European Resolution Infrastructure should be applied for resolving purposes and for identifier referral.

## *Short Term Agenda Issues for 2008 / 2*

### ■ (4) Metadata / Packaging Standard (complex!)

- Domain-specific **Dublin Core Application Profiles** to be developed and based on existing descriptive metadata standards to provide object-level search and retrieval across digital collections from libraries, museums, archives, institutional repositories, (inter-) national portals and other cultural heritage organisations.
- Each domain-specific Dublin Core Application Profile must include provision for **rights metadata** as well as some provision for **technical metadata** (at least the file format and the version of this format).
- For the provision of **collection level descriptive metadata** existing collection description formats (e.g. Michael, TEL, Archival Grid etc.) should be harmonised for use in the EuDL.
- A **Metadata Registry** for EuDL should be developed.

## *Short Term Agenda Issues for 2008 / 3*

### ■ (4) Metadata / Packaging Standard (continued)

- A higher level interoperability application profile should not be created. Instead, **semantic interoperability techniques** should be employed to implement **semantic mappings** and the cross-searching of descriptive metadata.
- **Packaging standards** such as METS, MPEG 21 (DIDL) or XFDU, that serve as “wrappers” for complex objects, should be considered as part of Issue 2 (Object Models).
- Section 5.1 of the Minerva Technical Guidelines can be used as a starting point regarding **file formats**. The work being done on file formats as part of the Planets project also needs to be considered.

## *Short Term Agenda Issues for 2008 / 4*

- **(5) Service Description Framework for Service Registry**

A [service registry](#) will be needed as part of EuDL; the JISC IESR repository could be a strong candidate as a starting point.

- **(6) Licensing Policies**

All freely available content and metadata should fall under a [suitable licence](#) clearly specifying the respective rights and use conditions.

- **(7) Authentication Data Exchange**

[Shibboleth-enabled methods](#) such as eduGAIN should be used as the standard solution for trust based exchange of authentication data within EuDL and towards the outside. A "What Federation Are You From" (WFAYF) service should thus be implemented as part of EuDL.

## *Short Term Agenda Issues for 2008 / 4*

- **(8) Basic Semantic Interoperability**

Make existing metadata and the **controlled terminology** used therein machine understandable to create a data layer ready for semantic query methods. The method of choice for conversion is SKOS, but use of OWL or RDF may be appropriate in some application scenarios.

- **(9) Awareness Building regarding Semantic Interoperability**

Demonstrate the added value to be gained from semantic interoperability and the **short term viability of converting existing controlled terminology** in experimentation environments relevant to the EuDL. These environments also to be used to market semantic interoperability functions of EuDL as **our unique selling point**.

- **(10) Interoperation of EUDL and WWW services (Google etc.)**

EuDL architecture should allow the creation of **maximum exposure** of services and content in generic WWW services (such as Google and Yahoo!) making sure the **EuDL provenance is clearly identifiable**.

## *Long Term Agenda Issues (2010 and beyond)*

- **Object Modelling** (increase in Granularity and Complexity)
- **Authorisation** (role models and role semantics),
- Usage **Logging, Accounting, Payment**
- Legal and **Access Protection** Issues (IPR / Rights / DRM)
- Advanced Semantic Interoperability (**Concepts / Ontologies / Rules / Reasoning**) and mapping to object modelling standards
- **Name Authority** Services
- **Multilingualism of Content**
- Identification of **Functional Primitives**
- **Service description** as a basis for service integration
- **Technical** and **Economical Sustainability**
- **Preservation Aspects**
- Strategic goal of EuDL: to act rather as a **service provider** or as a **data/object provider**?



- **Thematic continuity** is assured in that both Short Term and Long Term Agenda Issues are **input for EDLnet** and the WP2 Working Groups that will operate starting next week:
  - WG 2.1 Standards & Interoperability
  - WG 2.2 Semantic and Linguistic interoperability
  - WG 2.3 Technical Interoperability
- **Human continuity** is assured in that:
  - most EC WG members will be participating in the EDLnet working environment
  - I personally am leading EDLnet WP2 (Technical Interoperability) together with Makx Dekkers (DCMI)
- **Semantic Interoperability** and **semantic web integration** potentially distinguish EuDL from other web information services and may well constitute a strong unique selling point!
- Thank you for patience and attention!